Performance Evaluation For Modern Radars

Andrew G. Stove, David. L. Hurd,
Thales Sensors, Crawley, W. Sussex RH10 9PZ, U.K.

Abstract— **Modern radar systems are procured with tight specifications on a large number of different parameters. It is in the interests both of the customer and of the supplier that the procedures used evaluate radar performance are mathematically rigorous, precise and as cost-effective as possible. This paper describes some methods of evaluating the performance of different modes of modern radar systems and discusses the accuracy of which they are capable. The important place of modeling within these methods is emphasized.**

I. INTRODUCTION

Improvements in radar modeling are allowing designers to specify the performance levels very close to the theoretical limits. This leads to the very capable systems, but leaves little margin for the experimental uncertainty in evaluating their performance. The published literature on ways of evaluating radar performance is, however, surprisingly sparse. There are usually three separate phases in testing a new radar:

i) The first phase is to measure the parameters in the laboratory, to gain confidence that the radar will behave as predicted when it is taken 'into the field.'

ii) The second phase is generally the supplier's proving trials, which give confidence that the radar's behaviour is understood and hence that the formal acceptance trials will be successful.

iii) The third phase is the acceptance trials, witnessed by both supplier and customer, provides contractual evidence that the radar meets its specification.

The relationship between laboratory tests and field trials is discussed further in [1], and the importance of using an appropriate evaluation scheme for modern radar systems is discussed further in [2]. In this paper we shall refer to phase three as 'acceptance' trials, phase two as 'proving' trials and both together as 'evaluation' trials. This paper will concentrate on approaches which have been used by Thales, in co-operation with our customers, in these evaluation trials. Many of the experimental results described in this paper have been obtained during evaluation of variants of the 'Searchwater 2000' family of airborne surveillance radars[3], but the principles are also applicable to other types of radars.

The paper will look at evaluating three aspects of radar performance: noise-limited detection, clutter-limited detection in both non-coherent and coherent modes of operation and tracking. In addition to this paper, reference [4] discusses some of the issues involved in evaluating an automatic target classification system. The evaluation of imaging modes is a separate subject, which should draw on techniques used to evaluate photographic-type images, but taking account of the much greater dynamic ranges found in radar images.

The art of measuring individual parameters of radar equipment is a subject in their own right, but this paper discusses rather the principles involved in testing the top-level performance. In order to remain generic, actual performance values will not be mentioned, although the accuracies which they can be measured are discussed in quantitative terms.

Evaluation of a radar requires it to operate in a representative environment. Particular care must be taken in planning this for an airborne radar, because of the cost of installing the radar in an aircraft and flying it. The proving trials are generally less formal than the acceptance phase and may be carried out in a less-capable platform than that for which the radar is designed, which may then also be used to evaluate as much of the performance as possible, leaving only those requirements which need additional platform capability to be verified on the customer’s platform. This is a cost-effective approach which allows any problems with meeting the requirements to be addressed earlier, and reduces the number of flights required on the customer’s platform, which are likely to be more costly. This approach is in tune with the trend for ‘progressive acceptance’ of systems (i.e. incremental acceptance through life by validation of requirements as they become available). For Searchwater 2000, much use was made of a Douglas DC3 Dakota, which is cheap to fly and, being unpressurised, relatively cheap to modify.

II. NOISE-LIMITED DETECTION PERFORMANCE

The noise-limited detection is the simplest aspect of the performance to evaluate. It is the easiest case for which to calculate the theoretically-achievable performance, but the evaluation of the noise-limited performance uses many of the techniques which are also used in more complex scenarios, so this process will be considered in some depth.

Two methods have been used by Thales to assess the noise limited performance. The simpler method is to estimate the detection range against a known target from the fall-off in blip- to-scan ratio with increasing range. The other method is to measure the signal to noise ratio in recorded data.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **14 APR 2005**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Performance Evaluation For Modern Radars**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Thales Sensors, Crawley, W. Sussex RH10 9PZ, U.K.**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001798, Proceedings of the International Conference on Radar (RADAR 2003) Held in Adelaide, Australia on 3-5 September 2003.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**7**	

A. Blip-to-Scan Ratio

Measurements of the blip-to-scan ratio go back to the early days of radar: the target is observed for a number of scans of the radar and the proportion of scans on which it is detected, i.e. on which a 'blip' is observed, is calculated. Two issues with this method are the detection probability at which the range should be measured and how to smooth of the data.

Since the variation of detection probability with range is flatter for very high and very low probabilities, it is best to estimate the probability near the 50% point. It is also shown below that smoothing can bias the estimate at other probabilities.

In order to estimate the probability, it is clearly necessary to average the data over several scans. If the range, and hence the detection probability, does not change significantly between measurements, smoothing will reduce their variance. If, however the true probability changes from scan to scan, excessive smoothing can lose significant features and may bias the results. Fig. 1 shows a theoretical detection curve for a typical airborne target with 1, 7 and 17 point moving-average smoothing of the data. The solid, dashed and dotted lines correspond to the succesively higher degrees of smoothing. Note that the 50% detection point is almost unchanged by the smoothing, whereas the other points are moved. Smoothing using spline fits is sometimes preferred, since spreadsheet programs can do these automatically, but once again, one should make sure they do not bias the data.

Fig. 2 shows a noise-limited blip-to-scan curve from Searchwater 2000. The detection probability has been averaged over seven scans, which is our preferred length for air targets. The sharp fall-off with range allows the detection range to be estimated accurately. This sharp fall-off is typical of noise-limited performance with fast fading targets. The analysis of the blip-to-scan ratio has been used successfully in the proving and assessment of airborne radars.

The principal source of inaccuracy in this measurement is the uncertainty of the Radar Cross Section (RCS) of the target. Other, lesser, sources of error are fluctuations of the detections on individual runs and any unrecognized environmental effects, such as attenuation through intervening precipitation, rain clutter or surface clutter in the elevation sidelobes. For an airborne radar and an airborne target it is usually possible to arrange a geometry which eliminates both clutter at the range of the target and multipath.

In addition to the systematic errors due to imperfect knowledge of the mean RCS of the target and of the environment, a single run can be expected to show a standard deviation of about 13% in the detection range. The dominant contribution to this is the variation in the target RCS from run to run, which is expected to be about 2dB r.m.s., caused by small variations in the target aspect. This lead to an uncertainty of 12% r.m.s. in the detection range. To this is added about 4%r.m.s. due to statistical fluctuations between the measurements. The variance measured from a number of actual measurements was about 12%. The observed and expected errors are thus in good agreement.

B. Estimation the Signal to Noise Ratio

In order to estimate the signal to noise ratio seen against a trials target, the radar data must be recorded immediately after the analogue to digital converters. It is not necessary that the whole range swathe can be recorded, only that a sufficiently wide swathe is recorded to make it practical to record the region around the target.

Although indirect, the measurement of the signal to noise ratio can overcome the constraint of having to place an airborne target beyond the range of the clutter. The noise level can be measured in the absence of a target, preferably when the radar is not transmitting, and the signal can be measured separately, even if it is seen against a background of clutter. Provided that the linearity of the receiver is known, from laboratory tests, the signal to noise ratio can then be calculated. Relating the signal to noise ratio to the actual detection performance, however, requires auxiliary measurements to show that the data recording and the signal processing are performing as expected. The combination of the following measurements can be used to gain confidence in the process:

a) laboratory measurements of the components to predict the signal to noise ratio,

b) field measurements of the signal to noise ratio,

c) comparison of the radar display with the results of simulating the signal processor, to determine that the signal processing is behaving as expected, and

d) comparison of observed and predicted performance in partially clutter-limited conditions, which also require the noise-limited performance to be correct.

Since the signal level can be recorded against a clutter background, the measurements can be done against surface targets at relatively short range. For example a relatively small target of known RCS can be placed on a maritime platform, for example a Luneberg lens on a small ship, or else very large, static, corner reflectors have been used on land, being identified in the radar data by their size and by accurate knowledge of their position. The shorter ranges also make it much easier to characterize the environment over the whole radar path. The ability to use such targets allows this method to eliminate the uncertainty in the RCS of an airborne target, and in the environment over a long path.

Although indirect, this method thus overcomes the major systematic limitations inherent in blip-to-scan measurements, and has also been used successfully to evaluate airborne surveillance radars. Unlike the blip-to-scan ratio method, this

sort of technique was not used in the past because it requires a receiver with a high dynamic range and the ability accurately to locate the targets in the data, which is much easier now techniques based on GPS are readily available.

This technique can still be susceptible to multipath effects unless care is taken with the geometry. We have had most success by arranging geometries which avoid multipath, such as placing targets on cliff tops at short range, so that the multipath signal is only seen in the radar's sidelobes. Another approach, which we have found to be less successful, is to calculate the theoretical multipath gain and compensate for it, using calculations such as those described in [5], for example.

A examination of ten actual measurements from a single trial using a cliff-top reflector gave an r.m.s. uncertainty of about 2dB on a single measurement and a bias of also about 2dB.

III. DETECTION IN SEA CLUTTER

Blip-to-scan ratio measurements can also be made in clutter-limited situations, but determining a 'detection range' is generally inappropriate since the curve of detection probability against range is then often very flat. Fig. 3 shows such a curve for a maritime target seen by Searchwater 2000. Slow 'fading' of the target is seen, even with this frequency-agile radar, but there is no systematic reduction in performance with increasing range. Fig. 4 shows the same data averaged over 17 scans. This smooths out some of the fading but still fails to show a clear variation in performance with range. The additional dashed curve in Fig. 4, however, shows the predicted performance in the same scenario, indicating that the overall shape can be modeled. This allows a method of analysis which has been successfully used during proving trials: the detection probability is measured at a given range and then modeling is used to estimate the RCS required to obtain that performance in the trials environment. The RCS predicted by the model can then be compared with the known RCS of the target to see whether the sensitivity of the radar is as it should be. The standard deviation of a series of such measurements was estimated as about 1dB, although the absolute accuracy is again subject to the accuracy with which the RCS of the target is known.

The shapes of the detection curves in this case also mean that variations of the minimum detectable RCS are more meaningful measures than variations in detection range.

Performance estimates in clutter are of course subject to errors due to uncertainties in the characteristics of the clutter. As is well known [6], [7], this must be characterized not only by the mean level of backscatter but also its probability distribution and by its spatial and temporal correlations. An uncertainty of up to one 'sea state' is a reasonable estimate for the inaccuracy. This could lead to about 3dB r.m.s. uncertainty in the estimate of the performance in clutter-limited conditions.

If this uncertainty in the sea state is ignored, the variations in the target RCS still cause about 20% variation in the estimated detection range in clutter and the statistical fluctuations give another 10% variation. Practical measurements in clutter gave a standard deviation of about 23%. r.m.s., which is good agreement with this expected variation.

If different runs are not made in different directions, an additional source of systematic error is introduced by the effects of the swell and perhaps of the wind. It is usual for a company's proving trials to take place over a number or days, so these errors can be averaged out, but for formal acceptance trials this is not always possible. In some parts of the world data on sea conditions can be obtained from weather buoy data which is publicly available on the internet [8]. This data can be used to reduce, but not eliminate, uncertainties as to the characteristics of the clutter on a particular day.

A. Measurement of Clutter Parameters

The other approach to removing this uncertainty is to estimate the parameters of the clutter from data recorded during the trial, so that the performance estimates can use the actual clutter conditions rather than the nominal ones. For Searchwater 2000, the recorded radar data is also supplemented by recording key internal parameters of the signal processing. Reference [9] includes a discussion of the accuracy with which the required parameters can be estimated from the data. Instrumenting the radar is thus essential if meaningful assessments are to be made within the tight performance margins which are now often placed on radars.

B. Coherent Performance

In an ideal case, a coherent radar can completely separate the clutter from the target, so the performance becomes essentially noise-limited. This can happen with land clutter, but at sea some of the performance is often obtained from occasional detections in Doppler bins which contain small amounts of the clutter. This means that in order accurately to predict the performance, one must know the distribution of the clutter in the Doppler space. This, once again, can be obtained relatively easily if the clutter data is recorded during the trials.

Coherent operation allows the use as test targets of repeaters which modulate the signal, since the synthetic Doppler makes it easy to separate the repeater signal from the clutter and from the repeater's own 'skin return'. This technique has been used extensively to evaluate battlefield surveillance radars, such as that discussed in [4].

IV. CONNECTING MEASURED AND SPECIFIED PERFORMANCE

In the same way that it is sometimes impractical to obtain a real target with the characteristics called up in a radar's specification, it is frequently not possible, either, to find the specified clutter conditions. A way is needed to compare the trials results with the specification points. This must involve mathematical modeling. One approach is to use a model to

predict how a compliant system would behave in scenarios in which trials can be carried out. The customer can cross-check the supplier's modeling, using whatever models are available, but there is a risk that the supplier and customer will not be able to agree on the expected performance in the new scenarios. A more rigorous approach uses experimental results to validate the model under the actual trials conditions, and then uses the validated model to show that the radar would be compliant under the specified conditions.

V. STATISTICAL ANALYSIS

Proper statistical analysis is necessary to establish how much confidence the customer and the supplier can have in the results of any evaluation trials. The first stage in the trials planning is to decide what sort of experimental design is appropriate: if the aim is to establish the actual performance of the radar, as is often the case during proving trials, a two-sided test is required to place upper and lower limits on the uncertainty in the performance. Similarly, if the aim is to validate a model of the system, from which performance in various scenarios will be extrapolated, two-sided tests are again generally appropriate since a performance which is significantly better than predicted should cast doubt on the reliability of the model. If, however, the aim is directly to assess whether the radar meets its specification, then one-sided tests are normally appropriate. No-one will worry if the performance is better than specified.

The next step in the design is to estimate how many trials will be needed to obtain the required degree of confidence in the results. This requires information on their statistical nature. This plays an even more important role in validating models, when the statistical distribution predicted by the model should ideally be tested to see if it matches that observed in the trials.

To estimate the statistics, a model can be used with a Monte-Carlo process to generate an estimate of the distribution, with some limited accuracy due to the finite number of runs, or else appeal can be made to the central limit theorem and the errors can be assumed to be normally distributed. The standard deviation of the errors can also be estimated by using Monte-Carlo modeling , or else by using an *a priori* mathematical estimate. It is common practice to use *a priori* estimates of the errors, supported if possible by data from earlier trials, and then check once the trials results are available to see whether the assumed values were correct. Although this is a slightly dubious process, it is made more acceptable because inaccuracies in the estimates of the errors will only have a second-order effect on the trials results - they will not affect the actual results, only the confidence which the customer and the supplier can have in those results.

The process of assessment can be made more sophisticated by using the principle of *sequential testing* [10] to allow the trials to be prematurely curtailed if the radar can quickly be shown either to be clearly compliant or to be clearly not compliant. Limits on the performance are defined for each successive run, determined by the confidence required in the answers. If the average performance over all the runs exceeds the upper limit or falls below the lower limit, then the trials can be curtailed knowing that the compliance or otherwise of the system has been proved to a satisfactory level of confidence. Fig. 5 illustrates how the process works: the dotted curve shows the specification value; for a radar example this might be the required detection range, and the smooth curves show the limits. If the average performance exceeds the upper limit the system may confidently be said to be compliant, if it drops below the lower limit it may declared non-compliant. These limits converge on the nominal value as the number of trials increases. The jagged line shows a typical result, where the result of the first trial is below the nominal line, but not significantly so. In this illustration the average value crosses the limit at the fifth trial, so the sequence of trials can confidently be ended at this point. At worst, the process may continue until the original maximum number of trials has been completed, and then a simple pass/fail test will have to be performed, with the confidence actually achieved being calculated *post facto* from the trials data.

The number of trials required depends on three factors:

a) the uncertainty in the individual measurements

b) the minimum shortfall in performance to be detected

c) the risk of obtaining an erroneous result.

Obviously, the higher the uncertainties in the individual measurements the more runs are likely to be required to smooth out the random fluctuations. The further the actual performance is from the specified value, the more readily this can be noticed. Conversely, in the extreme case when an infinitesimally small deviation in performance must be detected, the number of trials would tend towards infinity.

The third significant factor is the chance of obtaining the wrong result: either falsely deciding that a non-compliant system is acceptable (the so called "buyer's" risk) or, alternatively, deciding falsely that a compliant system is not acceptable (the "seller's" risk). It is intuitively obvious that to lower the risk, more trials would be required. For an example case with 12% measurement uncertainty, 2dB detectable change in performance and 10% buyer's and seller's risks, the average number of trials would be about 5 and the test limits (the curves in Fig. 5) would be set at about $0.27R/N$, where R is the specified performance and N is the number of trials. The number of trials can vary dramatically, however. If the measurement uncertainty was 25%, and a non-compliance of 1dB must be detected with only 1% risk of an erroneous result, then on average 165 trials would be required.

The main disadvantage of sequential testing is the uncertainty in the number of runs which may be required. If the participants prefer to limit the number of runs, to limit their

commercial risk, then once the trials assets have been made available for that number of trials, the cost saving of then reducing the number of runs are often so minor that the participants decide it is better to carry out all the planned runs in any case. The case for sequential testing is stronger when individual runs are more expensive if, for example, each would requires firing (and destroying) a missile and its target. The maximum number of runs for which plans should be made would typically be about twice the average number required.

Sequential testing can also be applied to the two-sided problem of deciding whether or not a model is accurate and can be adapted to use the measured rather than the expected variance of the measurements[10].

VI. TRACKING PERFORMANCE

The second major aspect of radar performance which is evaluated in trials is tracking: at its simplest this involves comparing the variance of the tracker outputs with the specified limits. This can either be done by explicitly using an 'F' test [11], or a simplified version thereof, or else a simpler approach can be used by which a margin is added to the specified performance to allow for the expected uncertainties due to the limited number of trials and the uncertainties in the behaviour of practical targets. Calculations of these margins can use statistical procedures for comparing the expected and observed variances, but the allowance for deviations of the target trajectories must either be *ad-hoc* or based on a Monte-Carlo analysis of the effects of likely deviations. If actual target positions are not available, one must examine the data to check that their behaviour was within the expected limits. A single trial run may typically yield five independent sample points of the estimated track. Ten runs, yielding 50 samples, will then have a 95% probability of correctly detecting random errors which were 20% above the specified value.

It is now possible to instrument the target using differential GPS equipment, so that the radar's bias errors can be estimated, as well as the tracker's random errors. Basic statistical analysis can compare the observed and expected biases, taking account of the uncertainties introduced by the random errors. Assuming, as before, that there are 50 independent samples, there is a 95% chance of detecting bias errors which exceed the specification by more than about 25% of the standard deviation of the random errors.

Another approach is to look at the combined effect of the bias and random errors, which is compared with its expected value, taking account of the expected random errors, using a $\bullet^2$ test [12]. In this case, 50 samples give 95% probability of detecting a total error about 40% above the specification.

It will be appreciated that many of the same issues which were important in evaluating the detection performance also apply to the tracking performance. The essential first step in both cases is to determine the appropriate balance between confidence in the results and the number of trials which must be undertaken. Modeling may again be used to extrapolate from the specification points to the actual trials conditions. The need to allow for the actual behaviour of the targets can again be eliminated by feeding actual data into the model, as was recommended for detection performance evaluation.

Signal levels and clutter characteristics generally only have a second order effect on the behaviour of practical trackers. The false alarm rate, however, can have a major effect. Whereas an excessive false alarm rate helps the detection process, by allowing the radar to run at lower thresholds, of course it degrades tracking performance by allowing it to be seduced.

The way in which the tracking performance is specified and evaluated can significantly effect the effort required to conduct the trials: a specification of the tracking errors after a number of scans of tracking allows only one direct measurement for each run of the target, which requires a great number of runs to obtain accurate results. If the measured plot errors can be fed into a simulation of the tracker, however, more data can be made available. Alternatively, if the specification is, for example, the average error over a number of scans, more data can be obtained on each run, although care must be taken to ensure that the results are not confused by the correlation between adjacent tracker outputs due to its smoothing action.

One parameter of the tracker which is often omitted is the probability of losing a track: if the track operates at a very high probability of detection and a very low false alarm rate, the problems of track seduction and track loss can be avoided. However, in order to do this the radar must be running at a very high signal to clutter/noise ratio. The best compromise for a military radar, however, is often to initiate tracks at the lowest possible detection probability and the highest practicable false alarm rate, so the tracking performance is often specified under such conditions. There is then a significant probability that such a track will be lost. A complete tracker specification should therefore include a minimum probability of retaining a track under those conditions. A difficulty with such a specification may be that the engineers are not used to specifying this parameter, so there may be considerable uncertainty in knowing what sort of values are appropriate.

VII. CONCLUSIONS

The estimation of detection performance from blip to scan ratios still has a significant role to play in the evaluation of modern radars.

It is important to be able to extrapolate from the specification points to practical trials scenarios, so it is important to establish modeling results which can be agreed upon between customer and supplier.

It is important to be able to instrument the radar to gain good knowledge of the actual characteristics of the targets and the environment during the trials. Tracking performance can likewise be estimated more accurately if the target is fitted with a differential GPS system.

A complete specification of tracking performance requires appropriate consideration of the tracker's reliability.

The increasing complexity of the radars, and the increasingly stringent specifications which they have to meet, mean that a significant joint effort is needed by both the supplier and the customer to ensure that the assessment procedure is agreed well in advance of the commencement of the trials. The agreed procedure should be tested during the proving trials to ensure that it is actually workable.

The 'sequential' process of laboratory tests, proving trials and assessment trials is not really compatible with a 'concurrent engineering' approach to shortening development cycles. A step towards a 'concurrent ' methodology is initially to verify requirements by modeling, with the models being verified by trials on the supplier’s platform, and only a few trials being carried out at a later date on the 'target' platform.

VIII. ACKNOWLEDGEMENT

Many of the ideas described herein have been refined and developed in many discussions with colleagues within Thales and with customers' representatives.

REFERENCES

[1] S. Watts, "Specification and Measurement of Performance for Airborne Maritime Surveillance Radars," *Proc. 5th Int. Conf. on Radar Systems, Brest, 1999*, Session 2.6

[2] S. Watts, H . D. Griffiths, J. R. Holloway, A. M. Kinghorn, D. G. Money, D. J. Price, A. M. Whitehead, A. R. Moore, M. A. Wood and D. J. Bannister, "The Specification and Measurement of Radar Performance," *Proc. Radar 2002*, IEE Conf. Publ. 490, pp.542-546

[3] S. Watts and A. J. Gordon, "The Searchwater Family of Airborne Radar," *Proc IEE International Radar Conf, Radar 97*, IEE Conf. Publ. 449, pp.334-8.

[4] A. G. Stove and S. R. Sykes, "A Doppler-Based Automatic Target Classifier for Battlefield Surveillance," *Proc. Radar 2002,* IEE Conf. Publ. 490, pp.419-423

[5] Skolnik, M. I. "Introduction to Radar Systems," 2nd. Ed., 1980, New York, McGraw-Hill, Section 12.2.

[6] K. D. Ward, S. Watts and C. J. Baker, "Maritime Surveillance Radar Part 1; Radar Scattering from the Sea Surface," *Proc. IEE* vol. 137, Part F, pp.51-62, April 1990

[7] S. Watts, C. J. Baker and K. D. Ward, "Maritime Surveillance Radar Part 1; Detection Performance Prediction in Sea Clutter," *Proc. IEE* vol. 137 Part F, pp.63-72, April 1990

[8] http://www.noaa.gov/index.shtml

[9] S. Watts, "Modeling of Sea Clutter and Its Application to the Specification and Measurement of Radar Performance," *Proc 2001 CIE Int. Conf. on Radar*, pp. 431-5.

[10] O. L. Davies, "Design and Analysis of Industrial Experiments," London and Edinburgh, Oliver and Goldsmith,1960, Chapter 3.

[11] O. L. Davies, "Statistical Methods in Research and Production," 3rd Edition, London and Edinburgh, Oliver and Goldsmith, 1961, section 4.8.

[12] *ibid.* section 9.3

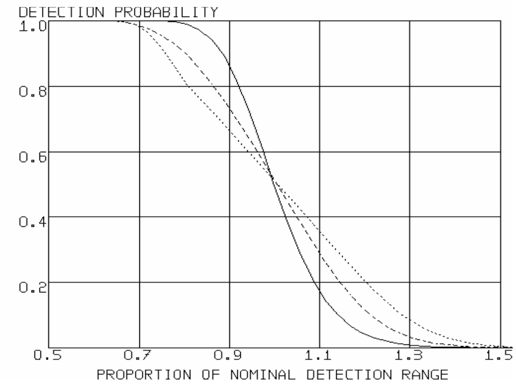

Figure 1: Effect of Different Levels of Smoothing on a Detection Curve

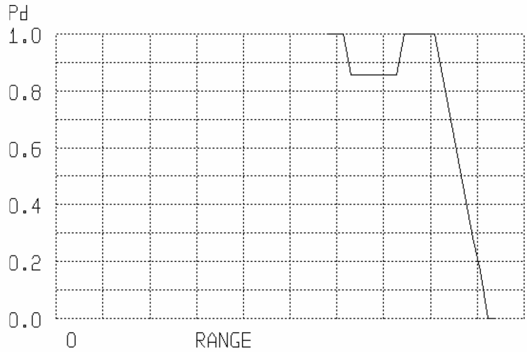

Figure 2: Blip to Scan Ratio curve for noise-limited performance, Averaged over 7 scans

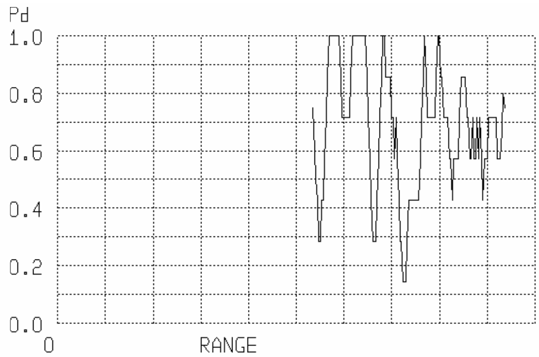

Figure 3: Blip to Scan Ratio curve for clutter-limited performance - Averaged over 7 scans

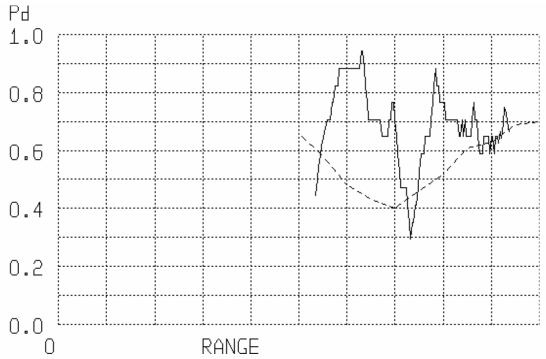

Figure 4: Blip to Scan Ratio curve for clutter-limited performance - Longer Average

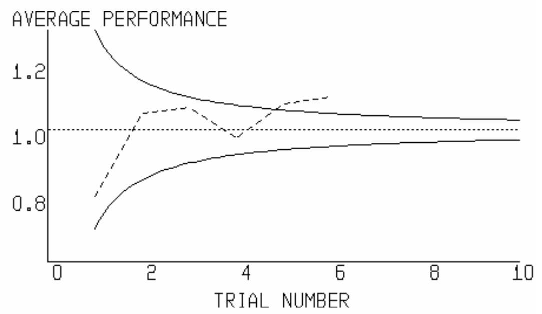

Figure 5: Illustration of a Sequential Test